JVIB

February 2004 • Volume 98 • Number 2

Research Report

A Computer System Serving as a Microswitch for Vocal Utterances of Persons with Multiple Disabilities: Two Case Evaluations

Giulio E. Lancioni, Nirbhay N. Singh, Mark F. O'Reilly, Doretta Oliva, and Gianluigi Montironi

The use of microswitches has been considered a crucial strategy to help individuals with extensive multiple disabilities overcome passivity and achieve control of environmental stimulation (Crawford & Schuster, 1993; Gutowski, 1996; Ko, McConachie, & Jolleff, 1998). In recent years, considerable efforts have been made to extend the evaluation of microswitches and to make their role more relevant. For example, studies have investigated (1) new, convenient responses and the matching microswitches that are likely to facilitate learning and performance with low-functioning persons and (2) the possibility of using multiple microswitches for multiple responses to extend motor activity and sensory input (Lancioni et al., 2002).

Responses to the environment by people with extensive multiple disabilities have almost exclusively involved simple motor acts, such as hand and head movements. However, vocal utterances may occasionally be present in a person's repertoire and could become crucial means of exerting control over environmental events (Lancioni, Singh, O'Reilly, & Oliva, 2003; Romski, Sevcik, & Adamson, 1999). To enable persons with severe multiple disabilities to reach the goal of using vocal utterances to control their environment, one would need microswitches that could discriminate between various utterances. Such microswitches do not exist. However, two computer systems were recently set up to substitute for such microswitches with one-syllable or wordlike utterances (Lancioni et al., in press).

This study evaluated the system's setup for wordlike utterances (ones that could have a wider application in daily contexts) with two participants with multiple disabilities. The system combines a new control software program with a commercially available speech- recognition program (Dragon Naturally Speaking, V6.0 Standard, manufactured by Scansoft, 2002).

Method

Participants

The participants (Janice and Bruce) were 20.5 and 18.9 years old. No IQ scores were available for them, but they were rated as being in the profoundly disabled range of intellectual ability. Janice had congenital hydrocephaly with secondary cerebropathy; Bruce had cerebropathy owing to prematurity and perinatal hypoxia. Both were blind and had spastic tetraparesis and sat in wheelchairs. They did not possess self-help skills, but presented satisfactory head and trunk control with useful hand and arm movements (for example, they could reach and touch or push objects on a table in front of them) and could emit some wordlike utterances. Janice had been involved in microswitch programs and continued to use them in her everyday life. Bruce had never used microswitches. They both lived at home with their parents and attended educational centers. Their parents provided formal consent for them to participate in this study.

Utterances and the computer system

Five and six utterances were used for Janice and Bruce, respectively. These utterances seemed the most comprehensible within the small group that the participants possessed and within those with more practical (plausible) connections with the participants' immediate environment (for example, a word such as *eyes* could be connected with songs concerning the eyes). Janice's utterances were her equivalents for *eyes*, *cough*, *grandma*, *talc*, and *go*. Bruce's utterances were his equivalents for *cat*, *sneeze*, *flute*, *guitar*, the name of a popular character, and the name of a song.

The system's hardware consisted of a Pentium III computer with a standard sound card, a screen, audio output devices, and a microphone for the participants. The system's software combined a new control program with a commercially available speech-recognition program (Dragon Naturally Speaking). The latter program did not allow for the reliable identification of the participants' utterances (given their poor quality). The control program (written with Borland Delphi 5, from the Inprise Corporation, 1999) allowed the association of each target utterance emitted by the participant with different words and phrases to be identified by the commercial software for that utterance over different trials. In practice, these words or phrases were grouped into specific categories through a variety of phonetic and component/length rules used by the control program. These categories then served as general references for the recognition of the target utterances and the delivery of favorite stimuli that matched them (discussed later). The control program also allowed for the occurrence of prompts if an utterance was not recognized (did not fall into a category), there was a break of one minute without utterances, or the same utterance was emitted in consecutive instances (for example, five times). The prompts consisted of various verbal encouragements to try again or to emit other utterances.

Favorite stimuli

Stimulus-preference screenings (Crawford & Schuster,

1993) identified favorite stimuli that were connected to the participants' target utterances (for example, grandma singing). The stimuli that were used for Janice's utterances included, among others, grandma talking and singing, music and songs, sharp noises, and air blowing. The stimuli that were used for Bruce's utterances included, among others, special songs, the sounds of cats and other animals, sounds of persons sneezing and coughing, and familiar people talking. The recognition of an utterance by the computer system produced a specific stimulus or set of stimuli that matched that utterance for 10 or 15 seconds.

Data collection

Three research assistants conducted the sessions (which lasted 15 and 20 minutes for the two participants, respectively) and recorded the participants' target utterances and whether the computer system recognized them correctly. Interrater agreement was assessed in over 15% of the sessions. The percentages of agreement that were computed on both measures (by dividing the number of agreements by the number of agreements plus disagreements and multiplying by 100) fell within the 80–100 range, with means exceeding 95.

Case evaluation

The evaluation included a baseline, an intervention, and a postintervention check. Prior to the intervention, there was a technical phase during which multiple recordings were made of the words or phrases identified by the computer system in relation to the target utterances emitted by the participants. These recordings allowed us to set up reference categories for such utterances (discussed earlier). To promote the participants' emission of these utterances, the research assistants used both modeling and favorite stimuli that were contingent on the utterances.

Baseline. No computer system or stimuli were available during the 10 and 8 baseline sessions conducted with the two participants. The research assistants modeled two or three of the target utterances at the start of the sessions. Modeling of the same utterances or the others that were available was repeated during the sessions if the participants did not emit any of the utterances for one to two minutes.

Intervention. The computer system was used during the 123 and 90 intervention sessions conducted with Janice and Bruce, respectively. The participants' utterances led the system to present matching stimuli (provided that it recognized the utterances correctly). Prompts would be given in the case of unrecognized utterances, emission breaks, and repetitions (discussed earlier).

Postintervention check. The postintervention check included 14 sessions, such as those implemented during the intervention, and occurred two months after the end of the intervention.

Results

The upper panel of Figure 1 shows the frequencies of occurrence of Janice's utterances, and the lower panel shows the frequencies of occurrence of Bruce's utterances. For each phase of the study, two bars are shown (each concerning half the sessions that were conducted) to depict possible changes across the phase. Janice's 10 baseline sessions showed a mean cumulative frequency for the target utterances (emitted independently of modeling by the research assistants) of fewer than 15 per session. During her 123 intervention sessions, the mean frequency of utterances was 34 per session. More than 70% of the utterances (with minimal variations between the first and second half of the sessions) were recognized correctly by the computer system and were followed by the matching stimuli. The other 25%–30% resulted in either a prompt to try again or the occurrence of nonmatching stimuli (that is, corresponding to another utterance mistaken by the system for the one emitted). A similar frequency was obtained during the 14 sessions of the postintervention check in which the mean percentage of recognized utterances exceeded 75.

Bruce's performance during the 8 baseline sessions was similar to that of Janice. His mean frequency of utterances per session increased to about 36 during the 90 intervention sessions and the 14 sessions of the postintervention check. The percentages of utterances recognized correctly were 68.5 during intervention and 71 during the postintervention check. No specific or strong trends were observed in the frequencies or percentages.

Discussion

The intervention outcome suggests that the computer system served as an adequate microswitch for the participants' vocal utterances. Although the technical phase may have had a large impact on increasing the participants' utterances early on in the study, the intervention and postintervention data underline the strength of the computer system in consolidating such responses at high levels and maintaining them at those levels over time. One could argue that the percentages of utterances that the system recognized correctly and that were followed by matching favorite stimuli were sufficiently high to motivate the participants to stay active. The procedures that were applied in cases of failures of recognition (that is, the prompts to try again and the occurrence of nonmatching stimuli) may also have played a positive role in this context (Miltenberger, 1997).

Improving the system's correct recognition of utterances is a critical objective for extending its use across persons and over larger sets of utterances (Hux, Rankin-Erickson, Manasse, & Lauritzen, 2000; Kotler & Tam, 2002; Venkatagiri, 2002). A first step toward this difficult objective may involve more extensive recordings of words and phrases that are identified by the computer system in relation to the target utterances. This step may allow one to build more comprehensive and accurate reference categories (Lancioni et al., in press; Wade, Petheram, & Cain, 2001). Additional software to control the loudness component of the utterances and the use of more sophisticated microphones may also be contemplated.

In sum, this study indicates that a computer system can serve as a microswitch for vocal utterances of people with multiple disabilities. The encouraging data need to be viewed with caution, given the preliminary status of the study and the lack of a definite experimental design. Future studies will have to extend the evidence and eventually improve the technology that is available.

References

Crawford, M. R., & Schuster, J. W. (1993). Using microswitches to teach toy use. *Journal of Developmental and Physical Disabilities, 5* , 349–368.

Gutowski, S. J. (1996). Response acquisition for music or beverages in adults with profound multiple handicaps. *Journal of Developmental and Physical Disabilities, 8* , 221–231.

Hux, K., Rankin-Erickson, J., Manasse, N., &

Lauritzen, E. (2000). Accuracy of three speech recognition systems: Case study of dysarthric speech. *Augmentative and Alternative Communication, 16* , 186–196.

Ko, M. L. B., McConachie, H., & Jolleff, N. (1998). Outcome of recommendations for augmentative communication in children. *Child: Care, Health and Development, 24*, 195–205.

Kotler, A.-L., & Tam, C. (2002). Effectiveness of using discrete utterance speech recognition software. *Augmentative and Alternative Communication, 18* , 137–146.

Lancioni, G. E., O'Reilly, M. F., Singh, N. N., Oliva, D., Piazzolla, G., Pirani, P., & Groeneweg, J. (2002). Evaluating the use of multiple microswitches and responses for children with multiple disabilities. *Journal of Intellectual Disability Research* , *46* , 346–351.

Lancioni, G. E., Singh, N. N., O'Reilly, M. F., Oliva, D. (2003). Extending microswitch- based programs for people with multiple disabilities: Use of words and choice opportunities. *Research in Developmental Disabilities, 24* , 139–148.

Lancioni, G. E., Singh, N. N., O'Reilly, M. F., Oliva, D., Montironi, G., Piazza, F., Ciavattini, E., & Bettarelli, F. (in press). Using computer systems as microswitches for vocal utterances of persons with multiple disabilities. *Research in Developmental Disabilities* .

Miltenberger, R. (1997). *Behavior modification: Principles and procedures* . New York: Brooks/Cole.

Romski, M. A., Sevcik, R. A., & Adamson, L. B. (1999). Communication patterns of youth with mental retardation with and without their speech-output communication devices. *American Journal on Mental Retardation, 104* , 249–259.

Venkatagiri, H. S. (2002). Speech recognition technology applications in communication disorders. *American Journal of Speech-Language Pathology, 11* , 323–332.

Wade, J., Petheram, B., & Cain, C. (2001). Voice recognition and aphasia: Can computers understand aphasic speech? *Disability and Rehabilitation, 23,* 604–613.

Giulio E. Lancioni, Ph.D., *Department of Psychology, University of Bari, Via Quintino Sella 268, 70100 Bari, Italy; e-mail: <g.lancioni@sico.uniba.it>.* ***Nirbhay N. Singh,*** *senior scientist, ONE Research Institute, P.O. Box 4657, Midlothian, VA 23112; e-mail: <nirbsingh52@aol.com>.* ***Mark F. O'Reilly, Ph.D.,*** *associate professor, Department of Special Education, University of Texas, Mail Code D5300, Austin, TX*

78712; e-mail: <markoreilly@mail.utexas.edu>. ***Doretta Oliva, M.A.,*** *research coordinator, Lega F. D'Oro Research Center, 60027 Osimo (AN), Italy; e-mail: <oliva.d@legadelfilodoro.it>.* ***Gianluigi Montironi, B.S.,*** *technical advisor, Lega F. D'Oro Research Center 60027 Osimo (AN), Italy; e-mail: <gianluigi-montironi@iseda.com>. Address all correspondence to Dr. Lancioni.*